Journal of Business and Economics, ISSN 2155-7950, USA
October 2014, Volume 5, No. 10, pp. 1754-1772
DOI: 10.15341/jbe(2155-7950)/10.05.2014/004

http://www.academicstar.us

The Implied Premium and Growth Strategy—Evidence from S&P 500

Sue-Fung Wang[1], Yi-Cheng Shih[2], Xiang-Jun Lai[1]
(1. National Chiao Tung University, Hsinchu, Taiwan; 2. National Taipei University, Taipei, Taiwan)

Abstract: Mispricing means the market price is deviated from intrinsic value. This difference, so called implied premium, might mainly be due to the information asymmetry from firm's growth strategy. The market tends to have a myth of paying too much premium for firm's growth. In general, firm pursue growth through diversification or focus strategy. Literatures have seldom discussed the relationship between implied premium and diversification. Therefore, the main purpose in this paper is to examine whether the implied premium is significantly associated with diversification or not. Our results show that diversified firms have higher implied premium. Meanwhile, the degree of diversification of high-tech firms has greater significant relationship than that of non high-tech firms. It implies that investors always pay too much for the growth of high-tech firms.

Key words: mispricing; residual income model; diversification
JEL codes: D82, G32, L25

1. Introduction

Classical finance theory argues that competition among rational investors, who diversify to optimize the statistical properties of their portfolios, will lead to an equilibrium in which prices equal the rationally discounted value of expected cash flows. Even if some investors are irrational, classical theory argues, their demands are offset by arbitrageurs and thus have no significant impact on prices.

Eleven years after the influential work of Fama (1970), formulating the efficient market hypothesis (EMH), Shiller (1981) criticized the EMH by providing empirical evidence on the so called overreaction hypothesis. He argued that price volatility is much higher than justified by changes in dividends which lead to periods of strong departures of stock market prices from fundamentally justified values. Among many others, DeBondt and Thaler (1985, 1987) and Chopra et al. (1992) present empirical studies supporting Shiller's hypothesis of prices overreacting to fundamentals. In contrast, Harris and Ohlsen (1990) and Bernard and Thomas (1989, 1990) find empirical evidence indicating that prices move less than fundamental information would justify. However, some supporters of the EMH do not believe that findings on over- and under-reaction are in contradiction to the EMH at all. According to Fama (1998), overreaction to information is equally likely as under reaction and both can be viewed as chance results. He argues that these "anomalies" disappear with methodical changes and that the

Sue-Fung Wang, Associate Professor, Graduate Institute of Finance, College of Management, National Chiao Tung University; research areas/interests: corporate finance. E-mail: sfwang@mail.nctu.edu.tw.
Yi-Cheng Shih, Assistant Professor, Department of Finance and Cooperative Management, College of Business, National Taipei University; research areas/interest: corporate finance and empirical investments. E-mail: ycshih.ntpu@gmail.com.
Xiang-Jun Lai, Graduate Institute of Finance, College of Management, National Chiao Tung University; research areas/interest: corporate finance. E-mail: saveyou250@gmail.com.

literature may present a biased sample of all studies developed. Since surprising results gain more attention and offer a larger possibility of being published, there may exist a bias towards the publication of such "anomalies".

Anomalies imply market price is deviated from intrinsic value[1] of a firm, i.e., there exist mispricing in capital market. Numerous literatures indicate that mispricing is resulted from investor sentiment (Chiang et al., 2011; Stambaugh et al., 2011; Baker & Wurgler, 2006). People from unreasonable expectation of likely returns and so make misguide consumption and investment decisions also easily make bubbles happened (Penman, 2010). Lakonishok et al. (1994) provides evidence that value strategies yield higher returns because these strategies exploit the suboptimal behavioral of the typical investor and not because these strategies are fundamentally riskier. In addition, information asymmetry may also cause mispricing. Information differences across investors (or groups of investors) have been a long-standing concern to price deviation. Uninformed investors require higher cost of capital due to less information obtained than informed investors. Information asymmetry issue can also be discussed among diversified and focused firms.

Firm value is created by investing and operating activities (Penman, 2010[2])[3]. And the firms' investing activities are associated with their growth strategies. We can discuss it on focus and diversification perspective by using three measures of diversification as the proxy variables for growth strategies[4]. Hyland and Diltz (2002) states that the typical firm diversified by making acquisition. Diversifying firms are poorly performing firms in comparison to specialized firms and have lower growth opportunities in their current activities. These diversifiers have accumulated a reserve of liquid assets. They can pay these back directly to shareholders, use the cash to diversify, or invest more in their current activities. The market anticipates that these firms will not return these liquid assets to shareholders and consequently may not be that surprised when firms make a diversifying acquisition. It might even be better for the firm to make such an acquisition that to use these liquid assets to finance investment in poorly performing operations. With this view, management diversifies to assure firm survival and growth when it faces difficulty competing within its industry. Each segment of a firm has its own investing strategy, and it may not consistent with the others. Therefore, it may generate information asymmetry among each segment.

Managers frequently cite the desire to mitigate asymmetric information as a motivation for increasing firm focus. An implication of this motivation is that diversified firms are subject to larger asymmetric information problems than are focused firms (Gilson et al., 2000; Habib et al., 1997). The source of the difference in asymmetry could be that diversified firms are less transparent than focused firms. Accounting figures for diversified firms are less transparent relative to those of focused firms. It is possible that asymmetric information problems are more severe for diversified firms. Aggregated cash flows and other diversification-related information problems make it more difficult for analysts (outsiders) to forecast firm cash flows as the transparency hypothesis. The transparency hypothesis predicts that, compared with focused firms, diversified firms should have, all else equal, larger forecast errors, more dispersion among analysts' forecasts, larger

---

[1] Intrinsic value is also known as true value, fair value, underlying value, or fundamental value.

[2] Stephen H. Penman (2010), *Financial Statement Analysis and Security Valuation* (4th ed.), International: The McGraw-Hill Companies.

[3] There are three main activities, operating, investing, and financing activities, of a firm. Operating activities try to maximize the profit of the firm by well operations. Investing activities use the cash raised from financing activities and generated in operations to acquire assets to be employed in operations. Both activities can create additional value for the firm. But the financing activities are investing activities for the claimants not for the firm. It can't add firm's value (Penman, 2010).

[4] The three measures of diversification are number of segments, Asset-based Herfindahl Index and Sale-based Herfindahl Index. The further information is listed in Section 2.1.2.

revaluations around earnings announcements, and smaller earnings response coefficients (ERCs). To the extent that they are less transparent than focused firms, diversified firms will face more difficulty in raising capital, less stock market liquidity, and, therefore, higher costs of capital (Thomas, 2002).

Above all, previous studies indicate that mispricing is associated with information asymmetry and investor sentiment, and also found that diversified firms have higher degree of information asymmetry than focused firms. Meanwhile, firm attempts to assure survival and pursue growth through diversification strategy. However, there are few literatures discuss the relation between mispricing and diversification proxy for growth strategy. Therefore, we mainly examine the relationship between implied premium and diversification. The main purpose in this paper is to verify that diversified firms have higher implied premium, resulting from the higher degree of information asymmetry for diversified firms than focus firms. We further investigate this relationship by group.

We group our samples by SIC code to examine how the growth strategy affects the implied premium in different groups. We classify the whole sample into two types, one is including financial institutions and another is without them. Furthermore, we classified each group into high-tech and non high-tech firms[5]. Most of literatures excluded financial institutions (F/I) from their samples because that the financial structure for financial industry is quite different to other industries. However, some literatures indicate that the market value of financial institutions is also significantly associated with diversification. Laeven and Levine (2007) find that there is a diversification discount: The market values of financial conglomerates that engage in multiple activities, e.g., lending and non-lending financial services, are lower than if those financial conglomerates were broken into financial intermediaries that specialize in the individual activities. While difficult to identify a single causal factor, the results are consistent with theories that stress intensified agency problems in financial conglomerates engaged in multiple activities and indicate that economies of scope are not sufficiently large to produce a diversification premium. Thus, in this paper, we particularly examine the relation between implied premium and growth strategy for F/I-included and F/I-excluded samples.

From the event of .com Bubble in the period of 1998 to 2000[6], also called Internet or Technology Bubble, we can observe that investors are always eager to pay more for high growth firms cause that its market value is deviated from fair value. High-tech firms have higher capital expenditure, so they must have greater growth potential and people might anticipate they have better performance in the future than that at the present. In the research of Bessiere and Elkemali (2011), they verify the hypothesis that if analysts exhibit overconfidence, they will overreact before the announcement and underreact after the announcement, and the misreactions (described former) will be greater for high-tech firms compared to low-tech firms is true. This hypothesis indicates that high-tech firms easily have greater misreactions for information. Hence, we divide our sample into high-tech and non high-tech firms to examine whether the implied premium of high-tech firms are really more significant associated with diversification than that of non high-tech firms.

The first problem in this paper is the way to determine implied premium. We refer Baginski and Wahlen (2003) but revising the formula slightly that we define implied premium as the difference between intrinsic value

[5] Following Brown et al. (2009), the largest three-digit high-tech industries are drugs (SIC 283), office and computing equipment (SIC 357), communications equipment (SIC 366), electronic components (SIC 367), scientific instruments (SIC 382), medical instruments (SIC 384), and software (SIC 737). The SIC code of financial institutions is 6000-6999.

[6] Thedot-com bubble (also referred to as the Internet bubble and the Information Technology Bubble) was a speculative bubble covering roughly 1995-2000 (with a climax on March 10, 2000, with the NASDAQ peaking at 5132.52 in intraday trading before closing at 5048.62) during which stock markets in industrialized nations saw their equity value rise rapidly from growth in the Internet sector and related fields (from Wikipedia, the free encyclopedia).

and actual fiscal year-end close price, that is, $\mathrm{IP_t} = (\mathrm{P_t} - \mathrm{IV_t})/\mathrm{P_t}$, where IP_t denotes implied premium at time t, IV_t denotes share intrinsic value at time t, and P_t is actual fiscal year-end close price at time t[7]. When IP is positive, it represents that firm value is overpriced by market. On the opposite, firm value is underpriced by market when IP is negative. For IV, we apply the residual income model[8] which is the present value of future residual earning (RE)[9] (Ebrahimi & Sarikhani, 2011; Higgins, 2011; Penman, 2010).

The second problem is the forecast horizon of RE. According to Richardson and Tinaikar (2004), there exist detective links between historical and forecast data branches, which often produce similar results. Moreover, long-tem analyst earning forecasts into RE have been proven not to improve pricing performance significantly (Lo & Lys, 2001). Some, such as Frankel and Lee (1999), continue to use shorter forecast horizon with one- and two-year ahead analyst earnings forecasts; however, these still suffer from biases in forecasting errors. The limitations encountered by previous studies suggest the validity of using historical EPS over forecast EPS in this paper.

The third problem with intrinsic value concerns required return of equity, denoted as r. r must be estimated, and is often viewed exogenous. Yoo et al. (2004) use CAPM-derived ICOE because individual betas predict positive and symmetric association ICOE in the literature. Banginski and Wahlwn (2003) indicate that the accounting-related risk measurements (i.e., the systematic risk and total volatility in a firm’s time series of residual return of equity) are associated with the market’s assessment and pricing of equity risk. Furthermore, their results show that the explanatory power of total volatility is incremental to the Fama and French (1992) factors, market beta, firm size, and the market-to-book ratio. Hence, in our research, we follow Hahn and Lee (2009), using FFr_t as our required return of residual income model, and also denoted as r[10].

Several firm characteristics, including sales growth, financial constraints, growth opportunities, and growth of profitability, are viewed as minor independent variables. These variables are recognized as existing influences on firm value in extensive prior studies. We consider that investors may regard a firm with growth of sales, opportunity, and profitability as a firm with excellent profit performance in the future. And they will be willing to pay more for it, resulting in higher implied premium.

The remainder of this paper is organized as follows. Section 2 describe the methodology, including sample selection, research model, and determination of intrinsic value. Section 3 presents and discusses the empirical results. Section 4 provides a summary of our main findings and the conclusion.

---

7 In Baginski and Wahlen (2003), the definition of mispricing, $\mathrm{IP_t} = (\mathrm{IV_t} - \mathrm{P_t})/\mathrm{P_t}$.

8 According to Penman (2010), Firm’s intrinsic value derived from residual income (also known as residual earning) model is consist of three parts, book value, value from short-term forecast, and value from long-term value (continuing value or terminal premium). The detailed information is provided in Section 2.1.1.

9 Residual earning is also known as economic value added (EVA) or residual income (RI). Basic formula of RE in per share is listed below:

$$\mathrm{RE_t} = \mathrm{EPS_t} - \mathrm{r_t} \times \mathrm{BPS_{t-1}}$$

where EPS_t denotes forecasted EPS at time t-1; r denotes required return of equity; BPS_{t-1} denotes book value of equity per share at time t.

10 The equation of FFr is listed below:

$$\mathrm{FFr_t} = \mathrm{R_t} - \mathrm{r_{ft}} - \sum_{k=1}^{k=4} \beta_k \times \mathrm{F_{kt}}$$

where R_{it} is stock return of firm i at time t, r_{ft} is the risk-free rate at time t, and F_{kt} denotes one of the Fama an d French four-factor loading (MKT, SMB, HML, and MOM).

2. Methodology

2.1 Data Collection and Variable Definition

The sample consists of S&P 500 members (COMPUSTAT auto-selection in 2011) over the 1998 to 2007 period. Our sample of firms is the intersection of CRSP, COMPUSTAT, and WRDS, and Research databases meeting the following requirements, applied yearly from 2001 to 2010:

(1) From COMPUSTAT, we collect the financial statement and accounting data such as actual fiscal year-end close price (P_{it}), capital expenditure, book value of total assets, operating cycle, average payment period for accounts payable, debt ratio, and EBITDA.

(2) From CRSP[11], we estimate individual beta for at least 24-60 months ahead CRSP stock returns.

(3) From WRDS[12], we obtain segment information to compute the two measurements of diversification.

(4) From World Economic Outlook (WEO) published by IMF, we collect the data of GDP growth rate of U.S.

The definition of key variables is reported in Table 1.

Table 1 Definition of Key Variables

This table displays the definition of key variables. The sample period is from 1998 to 2007.

Measurement	Variables	Definition	Reference
Implied Premium	IP	$IP = (P_t - IV_t)/P_t$	Penman 2010; Baginski & Wahlen 2003
Growth Strategy	NoSeg	The number of segments	Duchin 2010
	S_HI	$HI = 1/\sum(SegSalse/Sales)^2$	Bowen & Wiersema, 2005
	A_HI	$HI = 1/\sum(SegAssets/Assets)^2$	Bowen & Wiersema, 2005
Sales Growth	ΔSales	$(Sales_t\text{-}Sales_{t-1})/\ Sales_{t-1}$	Pajuste & Benjamin, 2005
Growth Opportunity	CAPEX	Capital expenditure/book value of total assets	Duchin, 2010
Size	LnTA	Natural logarithm of total assets	Gozzi et al., 2008
Profitability Growth	ΔROA	ROA = EBITDA/book value of total assets $\Delta ROA = (ROA_t\text{-}ROA_{t-1})/ROA_{t-1}$	Hahn & Lee, 2009 Martin & Francis, 2010

2.1.1 Implied Premium

Baginski and Wahlen (2003) define the price differentials ($PDIFF_{it}$) as RFV_{it} minus P_{it}, where RFV_{it} is intrinsic value computed by residual income model, and P_{it} is the price per share for firm i as of April 1 of each sample year for which we have analysts' earnings forecast data. For regression analysis, they take the form as $PDIFF_{it}/P_{it}$. But we revise the equation slightly that we define implied premium as the difference between intrinsic value and actual fiscal year-end close price, that is, $IP_t = (P_t - IV_t)/P_t$, where IP_t denotes implied premium at time t, IV_t denotes share intrinsic value at time t, and P_t is actual fiscal year-end close price at time t. When IP is positive, it represents that firm value is overpriced by market. On the opposite, firm value is underpriced by market when IP is negative.

(1) Intrinsic Value (IV)

Felthman and Ohlson (1995) model the relation between a firm value and accounting data concerning operating and financial activities. Book value equals value for financial activities, but they can differ for operating

[11] COUMSTAT only provides individual beta for recent five years, so we follow the COMPUSTAT procedure for beta estimation of individual firms to compute r. Detail is listed in Section 2.1.1.

[12] COUMSTAT only provides segment data for recent five years, so we obtain these data over the 1998 to 2007 period from Historical Segment of COMPUSTAT from WRDS.

activities. Firm value is assumed to equal the net present value of expected future dividends, and is shown, under clean surplus accounting, to also equal book value of expected future abnormal earnings (which equals accounting earnings minus an interest charge on opening book value. It demonstrates that the conclusions hinge on the extent to which the accounting is conservative as opposed to unbiased. Further, the absence/presence of growth in operating activities is relevant if, and only if, the accounting is conservative.

Afterward literatures develop other ways to evaluate firm's intrinsic value, such as multiples analysis, free cash flow discounted model, dividend valuation model, residual income model, and abnormal earning growth model. The multiples analysis is the easiest but inaccuracy because it considers fewer financial information of the firm, and it doesn't put future profitability into consideration. In addition, Preinreich (1938), Edwards and Bell (1961), Peasnell (1982), Ohlson (1995), and others show that the dividend valuation model is equivalent to the residual income valuation model. Hence, we follow previous literatures, using the residual income model as the approach to compute the intrinsic value.

The computation of intrinsic value from residual income model is composed of Book Value, Short-term Forecast Value, and Continuing Value[13]. The first portion, Book Value, is known for sure, and so firmly anchors the valuation. The second is based on forecast for two years ahead. These are typically made with some confidence, but with less assurance than the book value component. The value from these forecasts is the sum of the present value of the one- and two-year-head residual earnings. It forecast no growth in residual earnings after two years. The third portion adds value for growth. The long-term growth rate is usually fairly uncertain, so the component of the valuation is the most speculative (Penman, 2010). We estimate the growth rate as GDP growth rate of U.S after two-year-ahead[14]. The formula of residual income model to compute intrinsic value is shown below:

$$\mathrm{IV_t} = \mathrm{BPS_t} + \frac{\mathrm{RE_{t+1}}}{(1+\mathrm{r_t})} + \frac{\mathrm{RE_{t+2}}}{(1+\mathrm{r_t})^2} + \frac{\mathrm{RE_{t+2}}(1+\mathrm{g_t})}{(\mathrm{r_t}-\mathrm{g_t})(1+\mathrm{r_t})^2} \tag{1}$$

Where IV_t denotes share intrinsic value for firm i at time t, RE_t computed by EPS represents residual earnings per share of each firms at time t, and r represents a firm's required return rate[15]. g denotes firm's perpetual growth rate at a constant rate. We calculate intrinsic value by historical EPS (earnings per share), BPS (book value per share), and DPS (dividend per share) collected from COMPUSTAT[16].

The most important for residual income model is r and g. Because we can't obtain forecast growth rate for 3-5 year from I/B/E/S on Data stream, we follows Penman (2010) using the GDP growth rate as g[17]. We obtain GDP growth rate from World Economic Outlook (WEO) of IMF. But this may still suffer some bias. On the other hand, the determination of r is chosen from two types of measurements of required return of equity, reported completely in next section.

(2) Determination of Discount Rate (r)

---

[13] Continuing value can be measured under constant growth rate assumption or zero growth assumption. We take the former assumption to fit the reality as likely as possible.

[14] Penman (2010), states that GDP growth rate can be used as the perpetual growth rate. We obtain the data of GDP growth rate in U.S. from World Economic Outlook (WEO) published by IMF.

[15] r represents the required return of equity, derived from Fama-French four factors, which is listed completely in next section "2.Determination of r".

[16] In general, we compute the continuing value of residual income model by using forecast data of EPS, DPS from I/B/E/S on Datastream, and the forecast BPS which is determined by the following equation, $BPS_t = BPS_{t-1}+EPS_t-DPS_t$. Richardson and Tinaikar (2004) claim that there exist detective links between historical and forecast data branches, which often produce similar results. It is hard to obtain forecast data (lack of I/B/E/S database) so we use historical data to substitute it.

[17] Penman claims that firm's long-term growth rate must not be larger than national GDP growth rate in general. Thus, Penman use national GDP growth rate as the perpetual growth rate of firms.

We have considered two measurements as our r, CAPM-derived ICOE or Fama-French expected return (FFr). We choose FFr as r finally.

We follow the COMPUSTAT procedure for beta estimation of individual firms. The data is only traceable within five years of the present date. For data out of this range, we adapt the formula and steps set in the database using S&P 500 Index returns as market returns (R_M), risk-free rate (r_f), stock returns of each firms (R_i) in monthly data form, to estimate current individual beta (β_i). At least 24-60 previous observations are required to meet the regression requirements. We then substitute individual beta (β_i) into CAPM to obtain ICOE for individual firms. The formula of ICOE computation is presented below:

$$\text{CAPMderived ICOE}_t = r_{ft} + \beta_t(R_{Mt} - r_{ft}) \tag{2}$$

Where β_t is the beta of firm at time t, R_{Mt}-r_{ft} is the market premium at time t, and r_{ft} is the risk-free rate of U.S at time t[18].

We estimate the Fama-French expected stock return (FFr) by the procedures in (Hahn & Lee, 2009): estimating the Fama and French factor loadings (β_k) for individual stock i using monthly rolling regressions with a 60-month window every month requires at least 24 monthly return observations in a given window and substituting those betas into the model to obtain expected stock returns. The equation of computing FFr is reported as following:

$$\text{FFr}_t = R_t - r_{ft} - \sum_{k=1}^{k=4} \beta_k \times F_{kt} \tag{3}$$

Where R_t is stock return of firm at time t, r_{ft} is the risk-free rate at time t, and F_{kt} denotes one of the Fama an d French four-factor loading (MKT, SMB, HML, and MOM)[19].

Banginski and Wahlwn (2003) indicate that the accounting-related risk measurements (i.e., the systematic risk and total volatility in a firm's time series of residual return of equity) are associated with the market's assessment and pricing of equity risk. Furthermore, their results show that the explanatory power of total volatility is incremental to the Fama and French (1992) factors. Scholars also think that FFr can reflect the required return of equity better than CAPM-derived ICOE because it considers more risk factors. Thus, we measure firm's required return of equity by FFr, denoted as r, of residual income model, and the regression result of using FFr is significant, presented in Section 4.

2.1.2 Diversification Measures (Major Independent Variable)

Proxy variables of diversification are extensively discussed in many literatures. Conventional wisdom among finance scholars suggests that corporate diversification, especially conglomerate diversification, destroys shareholder wealth such that the shares of diversified firms sell at a discount. This link between diversification and value destruction is made in virtually every finance text. For example, a leading MBA finance texts put it this way, "diversification, by itself, cannot produce increases in value" (Ross et al., 1999). Furthermore, Brealey and Myers (2000) argue this is because "diversification is easier and cheaper for the stockholder than for the corporation." Yet, major U.S. corporations remain highly diversified. Montgomery (1994) identifies three main theoretical perspectives that can be used to explain why a firm might choose to diversify: agency theory, the resource based view, and market power.

Historically, corporate diversification has been measured using either the business count approach or the strategic approach. Following the business count method, diversification is assessed using Standard Industrial

[18] R_{Mt}-r_{ft} and r_{ft} are collected from Fama-French website.
[19] R_{it} is obtained from COUPSTAT. r_{ft} and Fama-French four-factor are all collected from Fama-French website.

Classification (SIC) codes[20] and corporate line-of-business data that are reported to the Securities and Exchange Commission annually. All of these measures share the common feature that they can be objectively calculated from publicly available data. The strategic approach is very subjective and relies less on SIC data and more on the judgment of the researcher (Martin & Sayrak, 2001).

The simplest business count measure of corporate diversification is the number of industry groups in which a firm operates. So we defined one of our diversification measurements as number of segment (NoSeg) collected from COMPUSTAT. COMPUSTAT only provides line-of-business information within recent five years, so we obtain these data from WRDS.

There's a problem with simply counting the number of SIC codes for the firm's different business units to measure diversification is that this measure fails to capture the relative importance or distribution of the firm's involvement in each industry segment. To solve this problem, Berry (1971) and McVey (1972) suggest that the use of the Herfindahl index, which was originally developed as a measure of industry concentration. The Herfindahl index can be used to capture the relative importance of the firm's different business segments for a single SIC classification level. There are two types of Herfindahl index, Asset-based and Sale-based. Asset-based Herfindahl index is computed as $\sum(\text{SegAsset/Assets})^2$, and the other, Sale-based Herfindahl index is determined as $\sum(\text{SegSales/Sales})^2$.

In our research, we defined the other two diversification measurement as Sale-based and Asset-based Herfindahl Index. Since lower values of the Herfindahl index indicate higher levels of diversification we instead use the inverse measure, $\text{HI} = 1/\sum(\text{SegSales/Sales})^2$ for consistency with the other diversification measures used here. This inverse measure equals one for a single business firm and it rises with the level of diversification (Bowen & Wiersema, 2005).

2.1.3 Minor Independent Variables

Firm value is created by operating and investing activities. We have mentioned that growth strategy is a kind of way of investing. In this paper, we mainly discuss the relationship between implied premium and growth strategy, which is proxied by diversification, and we consider it as our major independent variable in our regression model. In order to regress our model more completely, we further add some variable affect firm value as our minor independent variables in regression model. For institution, people prefer to invest firms with improvement of sales and operating performance. We use the percentage of change in sales as the proxy for sales growth (Pajuste & Benjamin, 2005). ROA[21] is considered as proxy variable for firm performance in a large body of literatures (Hahn & Lee, 2009; Mukherji & Pettus, 2008; Klapper, 2004). We use the change in ROA as the proxy for firm's improvement of operating performance (Martin & Francis, 2010). ROA can also be used to measure profitability. Firm make investment decisions to pursue growth opportunity, measured by the ratio of capital expenditure (CAPEX) to book value of total assets (Duchin, 2010). Previous studies usually classify state the growth opportunity of investment is realized as future profitability. Thus, profitability and investments are classified as categories of profitability as well (Hahn & Lee, 2009; Tim & Vidhan, 2008). As Penman (2010) says, "Don't pay too much for the growth." Growth of profitability has positive contribution to firms value (Hahn &

[20] SIC data is comprised of a four-digit scheme that can be used to define increasingly more refined measures of business or industry affiliation. The first two digits of the four-digit code "20" represent the broadest industry grouping. We take Food and Kindred Products for an example. After adding a third digit "201", we narrow the Food and Kindred Products group down to only those firms involved in Meat Products. Finally, adding a fourth digit "2013", we define the code for firms engaged in Food and Kindred Products—Meat Products—Sausages and other Prepared Meats

[21] ROA is the ratio of EBITDA to book value of total assets, and be interpreted as cash-based ROA (Aggarwal & Kyaw, 2006).

Lee, 2009), and as naturally we expect to observe a positive relation between investments, profitability, and implied premium.

We use the natural log if a firm's assets at the end of the year as the proxy for firm size (Gozzi et al., 2008)[22]. Firm size is considered a determinant of financial constraints or capital market excess (Timan & Wessels, 1988) that affects decisions of managers and firm value (Lee & Chuang, 2009). It is positively related to firm value (Maury, 2006) because small firms are younger and less well known, and there are therefore more likely to face financing constraints and vulnerable to capital market imperfections arising from information asymmetries and collateral constraints (Gertler & Gilchrist, 1994). Larger size firms have greater degree of information asymmetries, so resulting in higher mispricing (Thomas, 2002). However, in some cases, asset size also serves as proxy for firm risk. Some studies claim that size has positive effect on the risk taking of a firm due to the moral hazard associated with "too-big-to-fail" policy (Boyd et al., 2009), whereas others suggest a negative correlation between firm size and risk. Above all, we expect to find positive correlation between firm size and implied premium.

2.2 Sample Selection Criteria

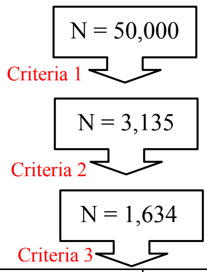
In addition, the sample selection criteria are as following:

(1) Criteria 1: missing data

(2) Criteria 2[23]: $r \leq 0$, $r \leq g$ and $g \leq 0$[24]

(3) Criteria 3: grouping by SIC code

We set four criteria for our sample selection. First, we exclude sample with incomplete data. Second, we retain samples satisfied with the assumption of residual income model. At the last, we group our sample by SIC code. The number of sample for each phase is listed in Figure 1.



	Whole Sample	High-tech firms	Non high-tech firms
With Financial institutions	1.634	367	1,267
Without Financial institutions	1.518	367	1,151

Figure 1 Number of Sample for Each Criterion

We have original data of 500 firms with 10 years (5,000 samples). After Criteria 1, we remain 3,135 samples due to trimming off samples with missing data. Then we obey the assumption of residual income model, remaining half approximately samples after Criteria 2. At the last, we group our final sample by SIC code. We can find that high-tech firms and financial institutions account for about 22.5% and 7.1% of whole sample respectively.

In the following, we attempt to develop the expected signs of coefficients of those variables reported in Table

22 Other proxy variables for firm size also exist, such as natural log of a firm's total sale or market value of equity.

23 According to Residual Income Model from Penman (2010), the required return of equity (r) has to be positive and more than growth rate (g), and the growth rate need to be positive. The detail of residual income model is listed in Section 2.1.1.

24 The growth rates are negative in 2008 and 2009, so the samples over the 2008 to 2009 period are excluded.

1in the following discussion. We provide descriptive statistics of key variables for whole sample in Table 2, by years and groups in Table 3 and Table 4, and we also provide Pearson Correlation Coefficients Test for whole sample in Table 5 and for groups in Table 6, all presented in Section 3.

2.3 Robust Regression Model

One of the most important statistical tools is a linear regression analysis for many fields. Nearly all regression analysis relies on the method of least squares for estimation of the parameters in the model. A problem that we often encountered in the application of the application of regression is the presence of an outlier or outliers in the data. Outliers can be generated by from a simple operational mistake to including small sample from a different population, and they make serious effects of statistical inference. Even one outlying observation can destroy least squares estimation, resulting in parameter estimates that do not provide useful information for the majority of the data. Robust Regression has been developed as an improvement to least squares estimation in the presence of outliers and to provide us information about what a valid observation is and whether this should be thrown out.[25]

There are two methods for robust regression, least squares alternatives and parametric alternatives, and we develop the former method by using SAS statistics software. For least squares alternatives method, the simplest methods of estimating parameters in a regression model that are less sensitive to outliers than the least squares estimates, is to use least absolute deviations. Even then, gross outliers can still have a considerable impact on the model, motivating research into even more robust approaches.

In 1973, Huber introduced M-estimation for regression. The M in M-estimation stands for "maximum likelihood type". The method is robust to outliers in the response variable, but turned out not to be resistant to outliers in the explanatory variables (leverage points). In fact, when there are outliers in the explanatory variables, the method has no advantage over least squares.

In the 1980s, several alternatives to M-estimation were proposed as attempts to overcome the lack of resistance. See the book by Rousseeuw[26] and Leroy for a very practical review. Least trimmed squares (LTS) is a viable alternative and is presently the preferred choice of Rousseeuw and Ryan[27]. The Theil-Sen estimator has a lower breakdown point than LTS but is statistically efficient and popular. Another proposed solution was S-estimation. This method finds a line that minimizes a robust estimate of the scale (from which the method gets the S in its name) of the residuals. This method is highly resistant to leverage points, and is robust to outliers in the response. However, this method was also found to be inefficient.

MM-estimation attempts to retain the robustness and resistance of S-estimation, while gaining the efficiency of M-estimation. The method proceeds by finding a highly robust and resistant S-estimate that minimizes an M-estimate of the scale of the residuals (the first M in the method's name). The estimated scale is then held constant whilst a close-by M-estimate of the parameters is located (the second M).

We conduct robust regression for the following regression equations:

Model 1: $IP_i = a_{i1} + \beta_{i1} \times NoSeg + \sum_{k=1}^{k=4} \beta_{ik,1} X_{ik}$

Model 2: $IP_i = a_{i2} + \beta_{i2} \times S\_HI + \sum_{k=1}^{k=4} \beta_{ik,2} X_{ik}$

Model 3: $IP_i = a_{i3} + \beta_{i3} \times A\_HI + \sum_{k=1}^{k=4} \beta_{ik,3} X_{ik}$

25 The description of robust regression analysis is cited from "Robust Regression" written by Lalmohan Bhar. And the following information about robust regression methods is obtained from Wikipedia, the free encyclopedia.

26 Rousseeuw P. J. and Leroy A. M. (2003), *Robust Regression and Outlier Detection* (1st ed.), John Wiley & Sons Inc.

27 Ryan Thomas P. (2008), *Modern Regression Methods* (2nd ed.), John Wiley & Sons Inc.

Where IP_i denote the implied premium for firm I, X_{ik} denotes minor independent variables as ΔSales, CAPEX, SIZE, ΔROA for firm i (definition of each variable is reported in Table 1). Our major independent variable, NoSeg, S_HI, and A_HI are three measures of diversification we've mentioned in Section 2.1.2.

3. Results

We first analyze our effective sample through descriptive statistics. Then we investigate the relationship between implied premium and diversification by robust regression. The results and analysis are presented below.

3.1 Descriptive Statistics of Key Variables[28]

Table 2 reports the summary statistics of key variables reported as mean and median for whole sample and by groups over the period from 1998 to 2007. In Table 2, we can find that IP are positive whether samples include or exclude financial institutions. It means that firm's market value is overpriced by market. Furthermore, mean and median of implied premium (IP) for high-tech firms is larger than that for non-high-tech firms, so we infer that the degree of misreaction of high-tech firms is larger than non high-tech firms. It is consistent with Bessiere and Elkemali (2011) that the misreactions will be greater for high-tech firms compared to low-tech firms. Besides, we can also find that financial institutions have positive IP, and it is less than that high-tech firms have, but larger than that non high-tech firms have. There is slightly difference of mean and median by group between F/I-included and F/I-excluded samples for diversification measured by NoSeg, S_HI and A_HI. Non high-tech firms have greater degree of diversification than high-tech firms have, but approximately equal to that financial institutions have.

Table 2 Descriptive Statistics of Key Variables

This table displays the summary statistics of key variables reported as mean and median for whole sample and by groups over the period from 1998 to 2007. Panel A shows the descriptive statistics, mean and median for sample including financial institutions and for high-tech and non high-tech firms. Otherwise, Panel B reports the descriptive statistics, mean and median for sample excluding financial institutions and for financial institutions and non high-tech firms. The number of sample is presented below, denoted as N.

Panel A. Samples include financial institutions						
	Whole sample		High-tech firms		Non high-tech firms	
	Mean	Median	Mean	Median	Mean	Median
IP(%)	41.02	82.58	57.30	91.29	36.30	79.70
NoSeg	3.21	3.00	2.38	1.00	3.45	3.00
S_HI	6.16	1.52	1.91	1.29	7.40	1.59
A_HI	8.84	1.77	2.73	1.37	10.61	1.86
N		1,634		367		1,267
Panel B. Samples without financial institutions						
	Whole sample		Financial Institutions		Non high-tech firms	
	Mean	Median	Mean	Median	Mean	Median
IP(%)	40.88	82.49	42.79	84.01	35.64	79.21
NoSeg	3.19	3.00	3.48	3.00	3.45	3.00
S_HI	6.47	1.52	2.14	1.46	7.93	1.60
A_HI	8.48	1.80	13.65	1.56	10.31	1.90
N		1,518		116		1,151

[28] We can infer the similar results from descriptive statistics by using risk-free rate as the implied cost of equity of residual income model to compute implied premium, denoted as IPrf. Mean and median of IPrf for high-tech firms is larger than that for non-high-tech firms. Financial institutions have positive IP, and it is less than that high-tech firms have, but larger than that non high-tech firms have.

Table 3 and Table 4 report the summary statistics of key variables reported as mean and median for whole sample by years. We can find that IP for each year are almost positive for each group in Table 3 and Table 4. In Table 3, we find that the degree of IP over the period from 1998 to 2000 is much greater than other years by whether the whole samples include or exclude financial institutions. It means that investors overestimated firm value due to too high anticipation for growth in that period. It also can be verified by .com bubble, which happened in the period of 1995 to 2000. People pay too much for the growth lead bubble happened, resulted in higher implied premium (Penman, 2010). It is consistent with Table 2 that the degree of misreaction of high-tech firms is larger than non high-tech firms by years Non high-tech firms have slightly greater degree of diversified than high-tech firms by years. There is no obvious change of mean and median of diversification measured by NoSeg, S_HI and A_HI for whole samples whether include or exclude financial institutions by each year.

We can observe the similar results as in Table 3 and Table 4. In Table 4, the degree of IP over the period from 1998 to 2000 is much greater than other years for high-tech and financial firms (listed in Panel A and D in Table 4). Besides, the degree of IP over the period from 1999 to 2000 is much greater than other years by whether the non high-tech firms include or exclude financial institutions (listed in Panel B and C in Table 4). These are consistent with .com bubble, which happened in the period of 1995 to 2000. There is no obvious change of mean and median of diversification measured by NoSeg, S_HI and A_HI for each gorup by each year.

Table 3 Descriptive Statistics of Key Variables for Whole Sample by Years

This table displays the summary statistics of key variables reported as mean and median for whole sample by years over the period from 1998 to 2007. Panel A shows the descriptive statistics, mean and median for sample including financial institutions. Otherwise, Panel B reports the descriptive statistics, mean and median for sample excluding financial institutions. The number of sample is presented below, denoted as N.

Panel A. Descriptive statistics of key variables for whole sample with financial institutions

	IP(%)		NoSeg		S_HI		A_HI		
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	N
1998	54.29	94.86	2.28	1.00	3.84	1.46	3.63	1.48	129
1999	77.33	98.59	2.92	3.00	2.63	1.57	2.81	1.71	83
2000	75.91	94.22	3.22	3.00	2.14	1.54	3.89	1.66	191
2001	48.59	85.94	3.06	3.00	1.95	1.45	4.27	1.66	172
2002	2.23	63.32	3.15	3.00	2.02	1.42	2.43	1.64	177
2003	12.47	75.72	3.14	3.00	1.87	1.40	4.28	1.70	170
2004	46.87	76.32	3.38	3.00	2.03	1.57	3.55	1.94	176
2005	46.85	83.30	3.33	3.00	1.93	1.47	3.90	1.85	175
2006	16.57	67.86	3.60	3.00	2.18	1.87	5.10	2.04	173
2007	51.60	82.72	3.63	3.50	2.20	1.79	3.58	2.02	188

Panel B. Descriptive statistics of key variables for whole sample without financial institutions

	IP(%)		NoSeg		S_HI		A_HI		
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	N
1998	53.45	94.86	2.26	1.00	2.39	1.46	2.87	1.70	123
1999	76.44	99.04	2.86	3.00	2.60	1.44	2.80	1.73	78
2000	74.54	94.07	3.24	3.00	2.15	1.54	3.04	1.65	178
2001	48.16	86.42	3.04	3.00	1.90	1.48	2.48	1.64	165
2002	-1.40	64.65	3.13	3.00	1.97	1.42	2.50	1.72	166
2003	17.33	76.00	3.12	3.00	1.88	1.43	2.59	1.99	158
2004	47.90	76.67	3.35	3.00	2.05	1.64	2.59	1.99	164
2005	41.92	82.58	3.28	3.00	1.91	1.47	3.00	1.92	153
2006	15.87	66.81	3.58	3.00	2.18	1.82	4.28	2.23	161
2007	52.18	82.94	3.65	4.00	2.24	1.85	3.21	2.06	172

Table 4 Descriptive Statistics of Key Variables for High-tech, Non High-tech Firms and Financial Institutions by Years

This table displays the summary statistics of key variables reported as mean and median by years over the period from 1998 to 2007. Panel A, B, C, and D report the descriptive statistics, mean and median for high-tech firms, non high-tech firms with financial institutions, non high-tech firms without financial institutions, and financial firms, respectively. The number of sample is presented below, denoted as N.

Panel A. Descriptive statistics of key variables for high-tech firms

	IP(%)		NoSeg		S_HI		A_HI		
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	N
1998	95.64	97.23	1.87	1.00	2.21	1.46	5.19	1.38	45
1999	100.71	100.16	2.17	1.00	3.08	1.37	3.16	1.52	36
2000	92.58	97.47	2.48	2.00	1.67	1.48	2.28	1.64	40
2001	58.72	91.71	2.36	1.00	1.65	1.00	2.03	1.08	50
2002	23.92	81.81	2.00	1.00	1.54	1.00	1.89	1.00	40
2003	44.38	78.88	2.24	1.00	1.56	1.29	2.02	1.47	41
2004	32.65	88.21	2.90	3.00	1.80	1.73	2.50	1.81	29
2005	73.94	88.38	2.65	2.00	1.74	1.15	3.00	1.42	26
2006	-27.33	65.42	2.59	2.00	1.84	1.15	2.55	1.92	27
2007	47.04	79.88	3.00	3.00	2.04	1.62	2.60	2.01	33

Panel B. Descriptive statistics of key variables for non high-tech firms with financial institutions

	IP(%)		NoSeg		S_HI		A_HI		
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	N
1998	32.14	91.14	2.50	2.00	1.50	1.45	1.08	1.64	84
1999	59.43	93.95	3.49	3.00	2.28	1.96	2.54	1.83	47
2000	71.50	93.53	3.42	3.00	2.27	1.55	2.27	1.55	151
2001	44.44	83.94	3.34	3.00	2.08	1.52	5.18	1.76	122
2002	-5.18	-1.76	3.49	3.00	2.16	1.56	2.58	1.77	137
2003	2.33	74.83	3.43	3.00	1.97	1.44	4.99	1.77	129
2004	49.68	74.38	3.47	3.00	2.08	1.50	3.75	2.00	147
2005	42.13	82.58	3.45	3.00	1.97	1.54	4.05	1.88	149
2006	24.69	69.41	3.78	4.00	2.24	1.96	5.58	2.09	146
2007	52.57	82.85	3.76	4.00	2.24	1.84	3.79	2.04	155

Panel C. Descriptive statistics of key variables for non high-tech firms without financial institutions

	IP(%)		NoSeg		S_HI		A_HI		
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	N
1998	29.11	89.34	2.49	2.00	2.65	1.45	2.70	1.65	78
1999	55.63	94.80	3.45	3.00	2.18	1.81	2.62	1.77	42
2000	69.31	93.21	3.46	3.00	2.29	1.58	2.96	1.81	138
2001	43.56	84.53	3.33	3.00	2.01	1.53	3.49	1.74	115
2002	-9.44	56.88	3.49	3.00	2.11	1.56	2.67	1.78	126
2003	7.85	75.57	3.43	3.00	2.00	1.47	2.67	1.77	117
2004	51.17	74.60	3.44	3.00	2.11	1.60	2.61	2.02	135
2005	35.37	81.62	3.41	3.00	1.94	1.47	3.00	1.98	127
2006	24.57	68.99	3.78	4.00	2.24	1.90	4.63	2.33	134
2007	53.40	83.02	3.81	4.00	2.28	1.91	3.35	2.16	139

Panel D. Descriptive statistics of key variables for financial institutions

	IP(%)		NoSeg		S_HI		A_HI		
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	N
1998	71.49	95.52	2.67	2.50	1.57	1.26	2.12	1.16	6
1999	91.35	90.21	3.80	4.00	3.14	2.66	1.88	2.13	5
2000	94.66	95.32	3.00	2.00	2.07	1.09	8.75	1.18	13
2001	58.80	65.45	3.57	3.00	3.14	1.28	3.04	1.82	7
2002	-14.82	37.78	3.45	4.00	2.75	1.33	1.62	1.41	11
2003	-51.46	33.24	3.42	3.50	1.74	1.37	7.65	1.51	12
2004	32.91	54.95	3.75	3.50	1.74	1.29	6.58	1.47	12
2005	81.12	90.83	3.68	4.00	2.09	1.95	5.11	1.63	22
2006	25.94	85.89	3.83	3.50	2.24	2.23	6.20	1.84	12
2007	45.35	82.45	3.38	3.00	1.82	1.36	7.56	1.71	16

Before we proceed with our discussion regarding expected signs of coefficients of variables, we must point out the collinearity problems exist in our model. We develop the Pearson Correlation Coefficients Test in Table 5 and Table 6 in the next section.

3.2 Discussion of Collinearity Problems

Pearson correlation coefficients test can examine the correlation between each variable. If independent variables have significant correlation with each other, the regression model might suffer serious collinearity problem, violating assumptions of regression. Thus, we take the Pearson test first, and then compute the VIF for each model by groups for further collinearity examination.

Table 5 reports the results of Pearson Correlation Coefficients Test for whole samples. In Table 5, we obtain similar results of Pearson Correlation Coefficients Test for whole samples whether include or excluded financial institutions. We observe that all independent variables are significantly related to dependent variable, IP, except for ΔROA. It is rational that NoSeg is positive related to CAPEX and LnTA. Typical firms diversified by making acquisition (increasing capital expenditure and firm size) to pursue growth (diversification) (Hyland and Diltz, 2002). It can be interpreted by the significant positively correlation between NoSeg and CAPEX (0.1646 with 1% significance level), and NoSeg and LnTA (0.4960 within 1% significance level) in Panel A, for instance. A_HI is also significantly related to LnTA (0.0607 with 5% significance level in Panel A, and 0.0612 within 5% significance level in Panel B).

Table 5 Pearson Correlation Coefficients Test for Whole Sample

This table displays the Pearson Correlation Coefficients Test of each variable for whole sample. Panel A shows the results for whole sample including financial intuitions. Panel B shows the results for whole sample excluding financial intuitions. IP is dependent variable in our research. Major independent variables are NoSeg, S_HI, and A_HI, and minor independent variables are ΔSales, CAPEX, LnTA, and ΔROA. The number of sample is presented below, denoted as N. *, **, and *** indicate significance level at 10%, 5%, and 1%, respectively.

Panel A. Whole sample with financial institutions samples (N = 1,634)								
	IP	NoSeg	S_HI	A_HI	ΔSales	CAPEX	LnTA	ΔROA
IP	1.0000							
NoSeg	0.1069 ***	1.0000						
S_HI	0.0041 *	-0.0205	1.0000					
A_HI	0.0100 *	0.0023	0.0359	1.0000				
ΔSales	0.0772 ***	-0.1020	0.0052	0.0084	1.0000			
CAPEX	0.0664 ***	0.1646 ***	-0.0132	-0.0323	0.0634 **	1.0000		
LnTA	0.1012 ***	0.4960 ***	0.0104	0.0607 **	0.1354 ***	0.1165 ***	1.0000	
ΔROA	0.0015	-0.0184	-0.0006	-0.0010	0.0038	-0.0079	-0.0120	1.0000
Panel B. Whole sample without financial institutions samples (N = 1,518)								
	IP	NoSeg	S_HI	A_HI	ΔSales	CAPEX	LnTA	ΔROA
IP	1.0000							
NoSeg	0.1014 ***	1.0000						
S_HI	0.0043 *	-0.0213	1.0000					
A_HI	0.0099 *	-0.0076	0.0359	1.0000				
ΔSales	0.0706 ***	-0.0884 ***	0.0055	0.0104	1.0000			
CAPEX	0.0683 ***	0.1629 ***	-0.0153	-0.0316	0.0646 **	1.0000		
LnTA	0.1090 ***	0.5033 ***	0.0126	0.0612 **	0.1321 ***	0.0576 **	1.0000	
ΔROA	0.0016	-0.0185	-0.0007	-0.0009	0.0039	-0.0094	-0.0117	1.0000

We take VIF test to illustrate the collinearity problem in Table 6. In statistics, the variance inflation factor (VIF) quantifies the severity of multicollinearity in an ordinary least squares regression analysis. It provides an index that measures how much the variance (the square of the estimate's standard deviation) of an estimated regression coefficient is increased because of collinearity. In practice, if VIF is larger than 10, collinearity problem exists; less than 10, the problem doesn't exist.

In Table 6, we observe that VIF are almost close to 1 for each model by group. Thus, we can conclude that there is no collinearity problem in our model.

Table 6 Variance Inflation Factor (VIF)

This table display the variance inflation factor (VIF) for each variable in each regression model. Panel A reports VIF for whole sample. Panel B and C reports VIF for high-tech and non high-tech firms. IP is dependent variable in our research. Major independent variables are NoSeg, S_HI, and A_HI, and minor independent variables are ΔSales, CAPEX, LnTA, and ΔROA. The number of sample is presented below, denoted as N.

Panel A. VIF for whole sample						
	with Financial Institutions Samples			without Financial Institutions Samples		
NoSeg	1.3488			1.3730		
S_HI		1.0003			1.0004	
A_HI			1.0047			1.0050
ΔSales	1.0223	1.0211	1.0214	1.0214	1.0212	1.0216
CAPEX	1.0317	1.0164	1.0169	1.0313	1.0070	1.0076
LnTA	1.3410	1.0312	1.0348	1.3548	1.0206	1.0243
ΔROA	1.0005	1.0002	1.0002	1.0005	1.0003	1.0003
N	1,634	1,634	1,634	1,518	1,518	1,518
Panel B. VIF for high-tech firms						
	with Financial Institutions Samples			without Financial Institutions Samples		
NoSeg	1.3067			1.3067		
S_HI		1.0104			1.0104	
A_HI			1.0033			1.0033
ΔSales	1.0712	1.0739	1.0708	1.0712	1.0739	1.0708
CAPEX	1.0093	1.0021	1.0019	1.0093	1.0021	1.0019
LnTA	1.3496	1.0776	1.0718	1.3496	1.0776	1.0718
ΔROA	1.0021	1.0009	1.0008	1.0021	1.0009	1.0008
N	367	367	367	367	367	367
Panel C. VIF for non high-tech firms						
	with Financial Institutions Samples			without Financial Institutions Samples		
NoSeg	1.3088			1.3476		
S_HI		1.0004			1.0006	
A_HI			1.0054			1.0060
ΔSales	1.0138	1.0132	1.0135	1.0155	1.0156	1.0160
CAPEX	1.0687	1.0462	1.0467	1.0702	1.0311	1.0317
LnTA	1.2921	1.0374	1.0412	1.2989	1.0179	1.0221
ΔROA	1.0019	1.0017	1.0017	1.0023	1.0020	1.0020
N	1,267	1,267	1,267	1,151	1,151	1,151

3.3 Robust Regression Results[29]

Robust Regression is used to eliminate the influence of outliers on regression results. The primary purpose of robust regression analysis is to fit a model which represents the information in the majority of the data.

We develop the robust regression by using three kinds of measures of diversification as the major independent variable to examine which measurement has the most significant association with implied premium. We also conduct our model by different groups. In panel A of Table 8, we observe similar results for whole samples whether include or exclude financial institutions. We find that diversification is positively correlated with IP, and NoSeg is the most significant. ΔSales is positively correlated with IP within 1% significance level. It means people are willing to pay more for firms with sales growth. CAPEX is positively correlated with IP within 1% significance level. It means people are willing to pay more for firms with greater growth opportunities. ΔROA is positively correlated with IP within 1% significance level. It means that people are willing to pay more for firms with profitability growth (growth of operating performance), but it is has lower significance level, 5%. LnTA is positively correlated with IP within 1% significance level. It is consistent with Thomas (2002), that firms with bigger size have larger degree of information asymmetry. Above all, firms with higher degree of diversification and growth of operating performance and sales, and larger size are easily overpriced by market. The adjusted R-square for whole sample whether includes or exclude financial institutions is located in 13-15%.

We further examine the robust regression by high-tech and non high-tech firms. We have mentioned that literatures state that high-tech firms have more growth potential than non high-tech firms. Investors pay high attention on high-tech firms' performance. Thus, the degree of misreaction of high-tech firms is larger than non high-tech firms. Compared with Panel C, we can find that for high-tech firms, all measures of diversification are much more significant positively correlated with IP than those for non high-tech firms. S_HI is the most significant. As we mentioned at the last paragraph, all minor dependent variables are significant positively correlated with IP. It means high-tech firms with higher degree of diversification and growth of operating performance and sales, and larger size are easily overpriced by market. For non high-tech firms, S_HI is not significant related to IP, but the other measures, NoSeg and A_HI are significant positively correlated with IP within 1% significance level. All minor dependent variables, except for ΔROA, are significant positively correlated with IP. Non high-tech firms have stable profitability due to fewer investing activities to pursue growth. Therefore, it is rational that ΔROA is not significant correlated with IP for non high-tech firms because there is no obvious change in ROA for non high-tech firms. The adjusted R-square for high-tech firms is approximately 18 or 19%, much higher than that of non high-tech firms, located in 10-14%.

[29] We also use IPrf as the dependent variable in the same regression. We can find that only NoSeg is significant correlated with IPrf, other two measures, S_HI and A_HI are not significant related with IPrf for each group. Besides, adjusted R-square is only 9-10%, lower than the regression we construct in context (13-15%) for whole sample whether includes or exclude financial institutions. For high-tech firms, the adjusted R-square is only 12-15%, lower than the regression we construct in context (18-19%). For Non high-tech firms, the adjusted R-square is only 4-6%, lower than the regression we construct in context (10-15%). Thus, we infer that IP is more effective than IPrf.

Table 7 Robust Regression

This table reports the regression coefficients from the robust regression model over the period from 1998 to 2007. It also reports the associated t-statistics in parentheses. Adjusted R-square for each model is also provided. Three measures of diversification are used as the main independent variable for each of three kinds of models. Panel A, B, and shows the results of robust regression for whole sample, high-tech and non high-tech firms, respectively. IP is dependent variable in our research. Major independent variables are NoSeg, S_HI, and A_HI, and minor independent variables are ΔSales, CAPEX, LnTA, and ΔROA. The number of sample is presented below, denoted as N. *, **, and *** indicate significance level at 10%, 5%, and 1%, respectively.

Panel A. Robust Regression for whole sample

	with Financial Institutions Samples						without Financial Institutions Samples					
Dependent variable: IP												
Intercept	104.40	***	108.01	***	108.37	***	108.59	***	112.95	***	113.36	***
	(30.46)		(31.8)		(31.92)		(29.98)		(31.41)		(31.47)	
NoSeg	1.5346	***					1.5597	***				
	(5.47)						(5.41)					
S_HI			0.0020	**					0.0018	*		
			(2.08)						(1.77)			
A_HI					0.0039	*					0.0042	*
					(1.7)						(1.82)	
ΔSales	0.1341	***	0.1444	***	0.1431	***	0.1125	***	0.1180	***	0.1167	***
	(8.15)		(8.48)		(8.43)		(6.95)		(7.03)		(6.97)	
CAPEX	0.2707	***	0. 3164	***	0.3222	***	0.3367	***	0.4008	***	0.4079	***
	(2.96)		(3.42)		(3.49)		(3.64)		(4.26)		(4.33)	
LnTA	2.4095	***	3.4282	***	3.4762	***	2.9364	***	4.0733	***	4.1321	***
	(5.85)		(9.29)		(3.69)		(6.58)		(10.2)		(10.33)	
ΔROA	0.0003	**	0.0003	**	0.0003	**	0.0003	**	0.0003	**	0.0003	**
	(2.43)		(2.29)		(2.3)		(2.46)		(2.29)		(2.28)	
Adj. R^2		0.1448		0.1268		0.1284		0.1545		0.1327		0.1343
N		1,448		1,448		1,448		1,345		1,347		1,347

Panel B. Robust Regression for high-tech firms

	with Financial Institutions Samples						without Financial Institutions Samples					
Dependent variable: IP												
Intercept	106.55	***	108.57	***	108.52	***	106.55	***	108.57	***	108.52	***
	(21.2)		(21.48)		(21.66)		(21.2)		(21.48)		(21.66)	
NoSeg	1.0446	**					1.0446	**				
	(2.18)						(2.18)					
S_HI			0.0656	***					0.0656	***		
			(2.71)						(2.71)			
A_HI					0.0914	*					0.0914	*
					(1.93)						(1.93)	
ΔSales	0.0640	***	0.0771	***	0.0757	***	0.0640	***	0.0771	***	0.0757	***
	(3.77)		(4.71)		(4.18)		(3.77)		(4.71)		(4.18)	
CAPEX	0.6217	***	0.6793	***	0.6789	***	0.6217	***	0.6793	***	0.6789	***
	(3.09)		(3.27)		(3.30)		(3.09)		(3.27)		(3.30)	
LnTA	2.1244		2.7928	***	2.7914	***	2.1244		2.7928	***	2.7914	***
	(0.00)		(4.7)		(4.75)		(0.00)		(4.7)		(4.75)	
ΔROA	0.0004	***	0.0004	***	0.0004	***	0.0004	***	0.0004	***	0.0004	***
	(4.08)		(3.82)		(3.85)		(4.08)		(3.82)		(3.85)	
Adj. R^2		0.1943		0.1831		0.1871		0.1943		0.1831		0.1871
N		321		324		323		321		324		323

Panel C. Robust Regression for non high-tech firms

	with Financial Institutions Samples						without Financial Institutions Samples					
Dependent variable: IP												

(Table 7 to be continued)

(Table 7 continued)

Intercept	100.80	***	103.83	***	104.40	***	106.51	***	110.22	***	110.78	***
	(23.23)		(23.81)		(30.46)		(22.53)		(23.83)		(23.93)	
NoSeg	1.5073	***					1.48895	***				
	(4.58)						(4.28)					
S_HI			0.0018						0.0015			
			(0.58)						(0.49)			
A_HI					1.5346	***					0.0044	***
					(5.47)						(1.84)	
ΔSales	0.1681	***	0.1739	***	0.1341	***	0.1336	***	0.1308	***	0.1290	***
	(7.18)		(7.22)		(8.15)		(5.6)		(5.47)		(5.41)	
CAPEX	0.2567	**	0.3140	***	0.2707	***	0.3664	***	0.4439	***	0.4525	***
	(2.43)		(2.93)		(2.96)		(3.35)		(4.12)		(4.19)	
LnTA	2.2561	***	3.2244	***	2.4095	***	3.0237	***	4.0492	***	4.1248	***
	(4.48)		(6.97)		(5.85)		(5.36)		(8.07)		(8.20)	
ΔROA	0.0000		0.0000		0.0003		0.0000		0.0000		0.0000	
	(-0.08)		(-0.01)		(2.43)		(-0.07)		(0.00)		(0.00)	
Adj. R^2		0.123		0.1043		0.1448		0.1302		0.1137		0.1162
N		1,119		1,124		1,123		1,020		1,019		1,019

4. Conclusion

Our results show the growth strategy affects the deviation between market value and intrinsic value (also so called implied premium) due to the information asymmetry. Generally, firms take either diversification or focus strategy to create firm value. Prior literatures indicate that mispricing is attributed to asymmetric information or investor sentiment. However, seldom literatures investigated the relationship between implied premium and growth strategy. Therefore, we contribute to the studies by examining this relationship. The evidence shows that diversified firms have higher implied premium, implying that investors overprice diversified firms. Furthermore, we examine the relationship between implied premium and diversification by industry. We find that the degree of diversification of high-tech firms have greater significant association to implied premium than that of non high-tech firms. It implies that market always overreact to the high-tech firms.

References:

Baginski S. P. and Wahlen J. M. (2003). “Residual income risk, intrinsic values, and share prices”, *Accounting Review*, pp. 327-351.

Baker M. and Wurgler J. (2006). “Investor sentiment and the cross-section of stock returns”, *The Journal of Finance*, Vol. 61, No. 4, pp. 1645-1680.

Berger P. G. and Ofek E. (1995). “Diversification’s effect on firm value”, *Journal of Financial Economics*, Vol. 37, No. 1, pp. 39-65.

Bessiere V. and Elkemali T. (2011). “Uncertainty and financial analysts’ overconfidence: European evidence between high-tech and low-tech firms”, *Working Paper*, University of Montpellier II.

Bowen H. P. and Wiersema M. F. (2005). “Foreign-based competition and corporate diversification strategy”, *Strategic Management Journal*, Vol. 26, No. 12, pp. 1153-1171.

Brown J. R. andFazzari S. M. et al. (2009). “Financing innovation and growth: Cash flow, external equity, and the 1990s R&D bom”, *The Journal of Finance*, Vol. 64, No. 1, pp. 151-185.

Chen F. and Jorgensen B. et al. (2004). “Implied cost of equity capital in earnings-based valuation: International evidence”, *Accounting and Business Research*, Vol. 34, No. 4, pp. 323-344.

Chiang M. C. and Tsai I. et al. (2011). “Fundamental indicators, bubbles in stock returns and investor sentiment”, *The Quarterly Review of Economics and Finance*, Vol. 51, No. 1, pp. 82-87.

Denis D. J. and Sibilkov V. (2010). “Financial constraints, investment, and the value of cash holdings”, *Review of Financial Studies*, Vol. 23, No. 1, pp. 247-269.

Duchin R. (2010). "Cash holdings and corporate diversification", *The Journal of Finance*, Vol. 65, No. 3, pp. 955-992.

Feltham G. A. and Ohlson J. A. (1995). "Valuation and clean surplus accounting for operating and financial activities", *Contemporary Accounting Research*, Vol. 11, No. 2, pp. 689-731.

Francis J. R. and Martin X. (2010). "Acquisition profitability and timely loss recognition", *Journal of Accounting and Economics*, Vol. 49, No. 1-2, pp. 161-178.

Gozzi J. C. and Levine R. et al. (2008). "Internationalization and the evolution of corporate valuation", *Journal of Financial Economics*, Vol. 88, No. 3, pp. 607-632.

Hahn J. and Lee H. (2009). "Financial constraints, debt capacity, and the cross-section of stock returns", *The Journal of Finance*, Vol. 64, No. 2, pp. 891-921.

Hughes J. S. and Liu J. (2007). "Information asymmetry, diversification, and cost of capital", *Accounting Review*, Vol. 82, No. 3, p. 705.

Hyland D. C. and Diltz J. D. (2002). "Why firms diversify: An empirical examination", *Financial Management*, pp. 51-81.

Lakonishok J., Shleifer A. and Vishny R. W. (1994). "Contrarian investment, extrapolation, and risk", *Journal of Finance*, Vol. 49, pp. 1541-1578.

Maury B. and Pajuste A. (2005). "Multiple large shareholders and firm value", *Journal of Banking & Finance*, Vol. 29, No. 7, pp. 1813-1834.

Ohlson J. A. (1995). "Earnings, book values, and dividends in equity valuation", *Contemporary Accounting Research*, Vol. 11, No. 2, pp. 661-687.

Rustichini A. (2011). "Decision making and equilibria", *Synthese*, pp. 1-12.

Stambaugh R. F. and Yu J. et al. (2011). "The short of it: Investor sentiment and anomalies", *Journal of Financial Economics*, Vol. 104, No. 2, pp. 288-302.

Stephen H. Penman (2010). *Financial Statement Analysis and Security Valuation* (4th ed.), International: The McGraw-Hill Companies.

Thomas S. (2002). "Firm diversification and asymmetric information: Evidence from analysts' forecasts and earnings announcements", *Journal of Financial Economics*, Vol. 64, No. 3, pp. 373-396.

Yalçın K. C. (2010). "Market rationality: Efficient market hypothesis versus market anomalies", *European Journal of Economic and Political Studies*, Vol. 3, No. 2, pp. 23-38.